Acoustics and Perception of Sopranos' Vowels: The Relationship Between Pitches and Vowel Formants and Its Effect on Vowel Intelligibility

Pik Yu Chan
The University of Hong Kong

1 Introduction

Operatic singers often aim to sing beautifully (to keep a consistent voice quality), loudly (to project their voices adequately), and remain understood (to keep clear diction) (Callaghan 2000). However, some vowels are harder to sing at higher pitches than others, and tend to be adjusted in higher pitches to achieve an even and smooth voice quality (Callaghan 2000). For example, the low F1 of high vowels make them harder to sing, thus a speech-like vowel space for [i] is often adjusted closer to [ɪ] when approaching higher notes (Miller, R. 2000). A speaker may speak in a pitch range of around 250 Hz and may pronounce [i] with a first formant of 440 Hz. Since the vowel [i] cannot be pronounced as such if the speaker sings at a pitch of 440 Hz, the speaker must adjust the F1 (vowel height) to be over 440 Hz. Imagine a singer who adjusts the vowel height of [u], raising the first formant. Since [u] has a lower F2 to begin with, the first and second formant becomes very close to each other. How then, can the singer differentiate [u] and [ɑ], which has a high F1 and low F2? Since F1 and F2 are adjacent to each other, the F2 (vowel backness) of [u] may need to be raised as well to make it clear that [u] is sung. The examples here suggest that if the pitch is changed, it is expected that vowel height and backness may need to be readjusted too. However, previous studies rarely discuss vowel changes in the soprano singing voice with reference to acoustic data. With the basic assumption that singers aim to maintain vowel integrity, we aim to explore how sopranos change their vowels across a wide range of pitches, and how such change affects vowel intelligibility. We will first introduce the concepts of timbre and register to understand how singers control F0. This is followed by an overview of the concept of "vowel modification" in vocal pedagogy. A production and perception experiment will then follow, with results showing that F2 may modify before F1, and that F1 may be a better predictor for changing vowel intelligibility across the pitch range. These findings raise questions regarding our existing understanding of the role of formants, drawing attention to the idea that formants are not only a mere linguistic parameter to determine vowel contrasts, but is also crucial to voice projection.

2 Background

2.1 *Timbre and Register* When two different singers sing the same music, we tend to like one singer's voice more than the other, because he/she sounds more beautiful. But what exactly are "beautiful" sounds? And what are the actual differences between the two voices? To answer these questions, two concepts are crucial: *timber* and *register*. Musically, the distinctive character or quality of someone's voice is known as *timbre* and can be described as the tonal quality of sounds (Campbell 2001). If two singers were to sing the same pitch with the same loudness, the wavelength and amplitude of the sound wave would be largely the same while the shape of the sound wave may differ (Rothenberg 1984). Human voices do not generate pure tones; instead, complex tones with many waves of different energy levels at varying frequencies are generated (Ladefoged & Johnson 2010). If a complex wave includes waves with less energy at higher overtones, the outcoming shape of the wave would be smoother and more rounded; if there were more energy at the higher overtones, the shape of the wave would look sharper (Rothenberg 1984). Such distinction can be understood as a role of timbre. The second relevant factor is *register*. Even if a person has a refined speaking voice, and sings well at low pitches, their voice may sound harsh or may even sing out of tune after passing a certain high pitch. This is because the human vocal range can be divided into different groups of adjacent pitches known as *registers* (Lewcock et al. 2001). When pitches are in the same register, they are sung in a similar manner; when singing pitches across registers, it is more difficult to sing them smoothly (Lewcock et al. 2001). When an untrained singer tries to sing through the registers, the sudden change in timbre or pitch jumps may occur, which is known as register breaks (Lewcock et al. 2001). The transitions

*Acknowledgements: Many thanks to Dr. Youngah Do for her guidance, comments and support throughout; Dr. Jonathan Havenhill for assistance in data collection and comments; Lai Ka Yau, Ryan for assistance in statistical analyses; and the six sopranos for their kindness, time, and beautiful voices.

Proceedings of *AJL 4*

between registers are known as the *passaggio*, and the *passaggio* range for each individual may differ (Callaghan 2000).

To understand singers' timbre and registers on a physiological level, their roles in actual voice production should be discussed. The voice source is generated at the larynx, specifically when air passes through the vocal folds, causing it to vibrate (Sundberg 1999). The sound wave that is generated includes multiple waves, and the wave with the lowest frequency (F0) determines the pitch of the sound (Sundberg 1999). There are also component waves occurring at every multiple of the fundamental, *harmonics* (Ladefoged & Johnson 2010). When the source passes through a vocal tract resonator, some of these component frequencies resonate with the vocal tract better as they form a standing wave inside the vocal tract, and are thus reinforced (Doscher 1994). Such resonant frequencies are known as formants (Doscher 1994). By constricting parts of our vocal tracts, formant frequencies can be controlled and manipulated. Producing the notes in a piece of music requires good control of the F0. To make the lyrics comprehensible, a singer would need to control F1 and F2 as well. Since the fundamental is adjacent to F1 and subsequently F2, changes in the fundamental may induce adjustments in F1 and F2. All of these formants must remain controlled when singing through registers such that timbral coherence is achieved. There is currently limited understanding regarding register changes, and most discussion still surrounds laryngeal mechanism. We ask if singers change their vowel space (F1 and F2) when singing through the *passaggio* and if so, what exactly is the nature of such change.

2.2 *Vowel Modification* Apart from singing with consistent timbre, operatic singers also often need to sing without amplification technology while keeping themselves heard over the orchestra (Doscher 1994). The strength in their voice is not controlled at the voice source, as straining the vocal folds with excessive pressure would be unsustainable for a lifelong career. A common solution to both keeping one's timbre consistent and to project one's voice is to rely on resonance (Doscher 1994). For sopranos, vowels are often adjusted at different pitches for better resonance (Callaghan 2000). This is known as vowel modification (Callaghan 2000).

Although vowel modification is used by most, if not all operatic singers, little is known about its acoustic properties. There have been attempts to outline how vowels change from one to another in the vowel modification process, but most of these studies have not been empirical. R. Miller (2000) describes vowel modification as a process where vowels subtly alter towards the schwa when the pitch rises. In the upper *passaggio* vowels are to be modified as follows: [i] to [ɪ]; [ɪ] to [e]; [e] to [ɛ]; [ɛ] to [ɑ]; [ɑ] to [ɔ]; [ɔ] to [o]; [o] to [ʊ]; and [ʊ] to [u]. However, after the upper *passaggio*, the vowels are not to be adjusted towards the schwa; instead, while the front vowels are modified towards the neutral center, the back vowels may require more closure (Miller, R. 2000).

Under Miller's model, vowel change appears to be systematic and gradual. Since this is not proven empirically, we ask what happens acoustically to the interplay between F1 and F2 at different pitch ranges if sopranos still aim to preserve vowel integrity.

We hypothesize that both F1 and F2 will change gradually as F0 changes in sopranos' singing. As the F0 changes gradually according to the melodic contour, F1 may be adjusted since F1 is closest to F0. Since vowel intelligibility depends on F1 and F2's relative distance, if F1 changes without F2 change, vowel integrity may be lost. Given F1 and F2's adjacency, it is predicted that F2 will also change gradually to preserve vowel integrity. Therefore, with the assumption that there will be a gradual change across F0, F1 and F2, F1 is predicted to change first, following F0 change due to their adjacency, and an F2 change will follow as a result of the F1 change.

This hypothesis will be tested through a voice production experiment. Acoustic data of sopranos singing different vowels across their vocal ranges will be collected and analyzed. The changes in vowel space at different pitches will be mapped with accordance to their formant changes. Based on where vowel modification starts on the F0 continuum, and whether it is located closely to the *passaggio*, data formalization will follow, with specific reference to the interplay between F0 and the vowel formants.

3 Experiment One

3.1 *Methods and Materials* Six trained sopranos participated in the experiment, all of them being native speakers of Cantonese. The sopranos were between 20 to 22 years old, and their years of training spanned from four to ten years. Five of them were pursuing a degree in music with voice as their main instrument, while the other has obtained the Associate Performance Diploma from Trinity College London (ATCL) in voice.

[a], [ɛ] and [ɔ] were included in the study, as they are common across Italian, French and German (*International Phonetic Association* 1999), which are common languages of western art music vocal repertoire. Since the vowels are also produced in Cantonese, stimuli could be presented in Chinese characters and participants could produce the intended vowels accordingly and naturally. [i] and [u] also satisfy such condition, yet when prevocalic conditions are concerned, they appear in complementary distribution in Cantonese (Kwan 2014). The vowels appeared in CV contexts, with the consonants [p], [t], [l]. All nine CV combinations were presented as tone one (high level tone) words. This results in the following list of stimuli: 巴 [paː˥], 啤 [pɛː˥], 波 [pɔː˥], 吋 [taː˥], 爹 [tɛː˥], 多 [tɔː˥], 啦 [laː˥], 哩 [lɛː˥] and 囉 [lɔː˥].

The CV combinations were sung across a range of around two octaves, from B3 (247 Hz) to C6 (1047 Hz), rising by semitone. The first and last note of the range were excluded from analysis. During the experiment, the soprano only needs to sing one consonant across the entire scalic passage, with the vowels sung in the order of [a] [ɛ] and [ɔ]. After all three vowels are sung at a given pitch, the soprano then sings the next pitch, which is either a semitone higher or lower. An audio file was provided to the participants to sing along to in order to keep pitch accuracy and to control vowel length. The tempo was set at 88 beats per minute. Words were presented as a musical manuscript, like lyrics.

The participants' spoken modal voice for all 9 CV combinations were also recorded as a separate task. The carrier phrase used was "佢話______啩" [kʰɵy˩˧ wa:˨ ____ kwa:˧], meaning "I guess he/she said____".

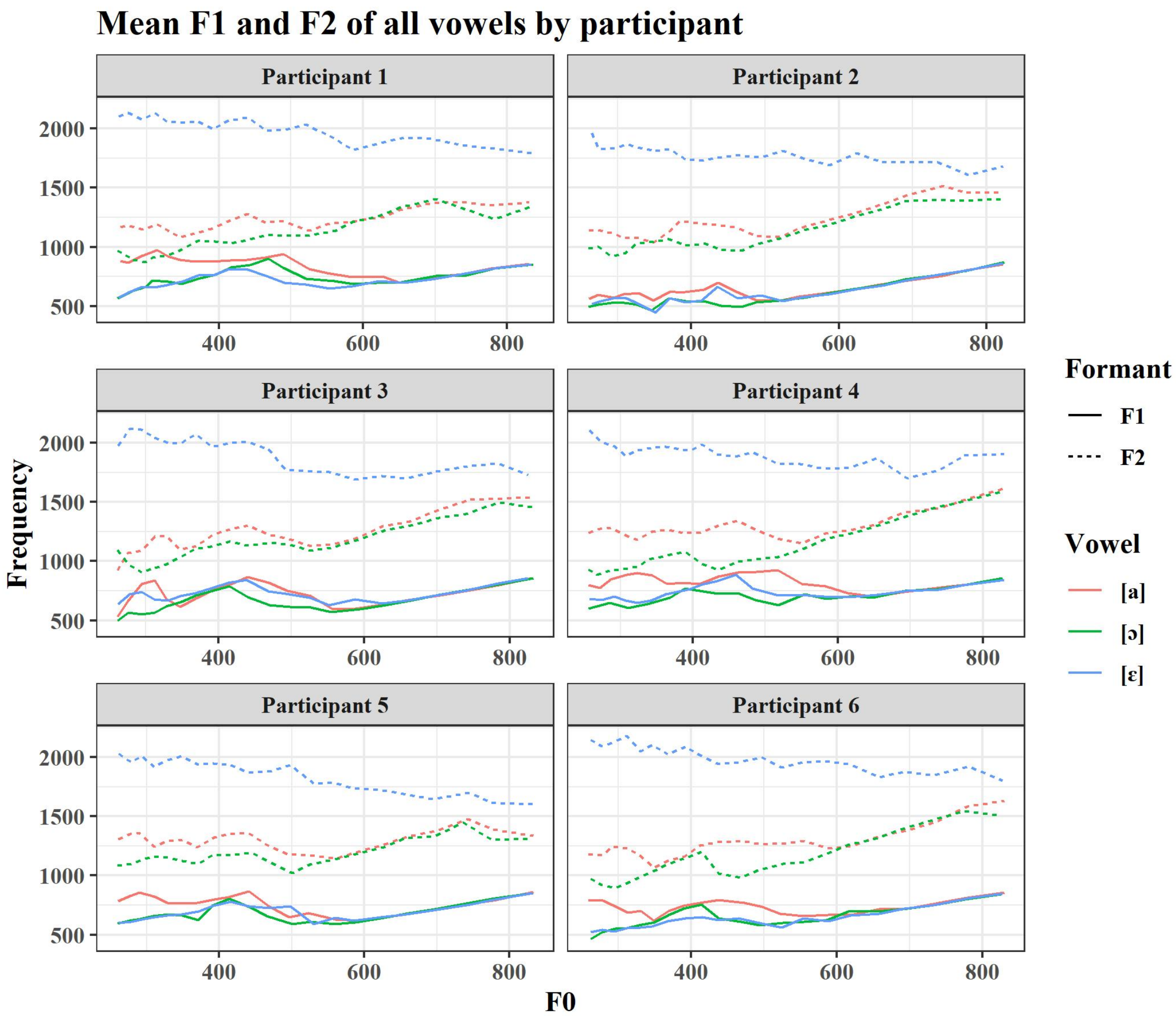
3.1.1 *Singing Task* Recordings took place at the sound booth of the Linguistics department at the University of Hong Kong. Two microphones were used, including a Sennheiser MD 46 microphone on a microphone stand and a Sennheiser mini clip on microphone. This ensures that even if the soprano's volume fluctuated across pitches, her voice would be picked up through her entire pitch range. The microphones were set at 44.1 kHz sample rate and 16-bit sample depth. Samsung Wired Stereo Earbuds were used to present audio stimuli, which was only used on one ear of the soprano's own choice, so that they can hear themselves during the task.

Participants were first given the visual stimuli presented on paper, and asked if they could read all the words out once, to ensure they know the intended pronunciation and tone. A practice trial using the stimulus set [pa:˥], [pɛ:˥], [pɔ:˥] was then sung once. The consonants' order of appearance was pseudo-randomized. Each scale was sung in an ascending-descending order, such that there are two values for each CV combination at each pitch, apart from 1047 Hz, which was excluded from data analysis.

3.1.2 *Speaking Task* The recordings were taken at the same location as above. The same Sennheiser clip on microphone was used alongside an Oktava mk 012 microphone. The microphones were also set at 44.1 kHz sample rate and 16-bit sample depth. Participants were seated at the computer, and the monitor was placed at eye level at the participant's most comfortable distance. The stimuli were presented in the carrier phrase in a randomized order, with two repetitions for each phrase. A 500 ms gap was included between each stimulus to indicate that the coming phrase is a new stimulus.

3.2 *Results* In total, 2862 tokens were collected, with 2700 tokens intended to be analyzed after excluding the recordings at B3 (247 Hz) and C6 (1047 Hz). The total number of tokens analyzed reduced in the end, as formants are only apparent when sufficient harmonics could reinforce them. When approaching the higher ranges, the harmonics inevitably spread out, and one can barely tell where the formants are. At the highest pitches, it is difficult to tell precisely where the intended formants are based on the spectrogram itself.

In the end, only 2376 tokens were analyzed, including all the tokens in modal (speaking) voice, and those sung from C4 (262 Hz) to G#5 (831 Hz). Notes above G#5 were eliminated. The formant values for F1 and F2 were taken at the nucleus of the vowel, but the mean F0 across the vowel was taken, in case of F0 fluctuations due to vibrato. Due to the small number of participants and individual variability, there will be no universal analysis across all participants, but instead, the results of each soprano are presented as follows. The x-axis shows the F0 frequency from low to high pitches, while the y-axis shows the frequency of the vowel formants in Hz.



It was predicted that both F1 and F2 will change gradually, and based on descriptive observation, it appears that both F1 and F2 for all vowels tend to change gradually, supporting the initial hypothesis. Interesting to note, is that across all participants, there seems to be fluctuations in both F1 and F2 measurements across all vowels at the lower half of the pitch range. This is surprising because the lower pitch range was expected to have similar and stable formant measurements compared to speech. At the upper half of the pitch range, vowel modification appears to become stable.

Looking at the specific formants separately, the F1 value tends to remain similar for most of the pitches. Only after the participant reaches the higher notes, does the F1 value lean closer to the F0, and eventually rise gradually in accordance to the fundamental. The point at which it starts modifying depends on the singer's general F1 range. Participants who have a higher formant range in general tends to modify their F1s later, at around E5 (659 Hz) or F5 (699 Hz) such as participants one, four and six, whereas the others modify the F1 a few notes earlier, at around C#5 (554 Hz) or D5 (587 Hz). It is also worth noting that the F1 values for all three vowels are generally close by, with the F1 for [a] generally staying above the F1 of [ɛ] and [ɔ] in the lower pitches, as expected. This is with the exception of participant three, whose F1 value of all three vowels tend to have more overlaps. The general trend of F1 supports the hypothesis for a gradual increase.

Expectedly, [ɛ] has the highest F2 value throughout and [ɔ] generally the lowest. From descriptive observation, it appears that all participants have a gradual F2 change for [ɛ] from the beginning. The F2 value tends to drop steadily across the pitch range. On the contrary, the F2 values for [a] and [ɔ], which are rather close by, both tend to modify after a certain pitch. Across all participants, there is a point at which the F2 values for [a] and [ɔ] converge and modifies in a similar manner thereafter. At that point, the vowels [a] and [ɔ] would theoretically be acoustically undistinguishable. Although the formant frequencies for the coming pitches could unfortunately not be mapped, it appears as if the F2 values for all three vowels will eventually converge. Since the general tendency is that F2 modifies gradually, the results appear to support our hypothesis.

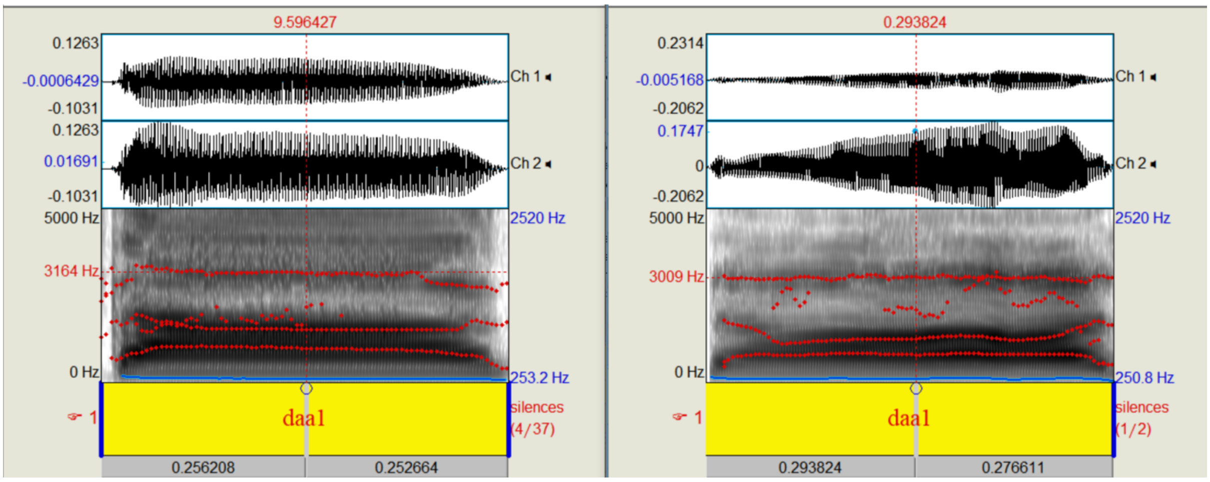
There are however, a few more unexpected details to note especially in the F2 patterns. First, the point at which F2 for [a] and [ɔ] starts modifying is worth noting. Across all six participants, F2 tends to start modifying between C5 (523 Hz) and D5 (587 Hz), even if the singer tends to modify F1 later. This suggests that instead of the F2 modifying because of the change in F1 as predicted, the location at which F2 modifies is independent of

F1, and the change in F2 is probably not a result of wanting to preserve vowel integrity, but for musical reasons instead. Moreover, after the F2 of [a] and [ɔ] converges, the F2s for both vowels tend to rise gradually up a slope very cleanly, without formant fluctuations.

3.3 *Discussion* The results do not demonstrate the same systematic direction of vowel modification towards the schwa as R. Miller (2000) describes. It was argued that [ɛ] would move closer to [ɑ] (lowered F2 and perhaps a slightly higher F1) and [ɔ] towards [o] (lowering F1), in the lower registers. After the second *passaggio*, the front vowels are to modify towards the neutral center (lowered F2), while the back vowels will require more closure (lowered F1) (R. Miller 2000). The results of this experiment suggest that this may not necessarily be the case for sopranos. [ɛ] generally has a lowering F2 and a stable or rising F1 across the participants and could be interpreted as moving closer to [ɑ]. However, the F1 does not tend to drop in the lower registers. Whereas, [ɔ] does not demonstrate a lowering F1, and may not be moving towards [o]. For vowels after the second *passaggio*, neither F1 nor F2 generally lowers itself. Only the F2 of [ɛ] drops. The results of this experiment provide evidence that female voices may require a separate model of vowel modification.

The experiment demonstrates vowel modification as a gradual process of formant change in accordance to a rising F0 in soprano voices. The results align with the claim that lower formants including F1 may take part in formant tuning in sopranos at higher pitches. Interesting to note, is how the results show a tendency of the F2 for [a] and [ɔ] getting closer to H2 at around C5 (523 Hz) to D5 (587 Hz). The results are supported by other studies that have also reported a tendency for the second formant to keep close to the second harmonic at the upper end of the middle register in the case of the vowel [ɑ] (Miller, D. 2000). Noteworthy in this experiment, is that the second formant aligns closer to H2 at around that range independent of F1. That is, the F1 may modify after F2 starts changing, and F2 is not a result of F1's modification. One possible reason for this is that the F2 may be used for reinforcing the second harmonic, causing the F2 to rise in a neat slope thereafter. It has been discussed in literature that formant-harmonic clusters at around 3000 Hz are crucial for singers to achieve a "vocal ring" (Doscher 1994). For musical notes lower in the vocal range, the harmonics may have been less wide, such that there may be stable harmonics at around 3000 Hz for a formant cluster to form. However, formant clusters may not be available for sopranos around C5 and onwards. For example, C5 (523 Hz) would have an H6 at 3138 Hz, C#5 (554 Hz) a H5 of 2770 Hz and H6 of 3324 Hz, then D5 (587 Hz) would have a H5 at 2935 Hz. The fluctuation of harmonics at that range suggests that there cannot be a stable singer's formant through the consecutive pitches thereafter. Thus, it could be that more reinforcing of harmonics using F2 would allow for better voice projection.

The prioritizing of voice projection over vowel integrity may also be supported by the difference between speech and singing. Since the fundamental frequency for speech does not differ largely from the first note of the scale (C4 262 Hz), it would be expected that the vowel formants should remain largely consistent in the lower range of the sung scale. However, when comparing the formant values for speech against the singing voice, the formants tend to be substantially higher in speech than in singing voice across all participants. Below is an example of [ta:˥] spoken (left) and sung (right) by participant four. The fundamental frequency for the spoken [ta:˥] is roughly around 253.2 Hz, while the F0 for the sung token was intended to be 262 Hz (C4); but in this token, it was slightly flattened to around 250.8 Hz, making it suitable for comparison. Looking at the formants roughly, for the spoken token, all of the first three formants are higher than that of a singing voice. Note also that the F3 for the sung token on the right seems to have been tuned to around 3000 Hz. This may be done via the lengthening of the vocal tract, which would lower all three formants. Although scholars have mainly been attributing the "singer's formant" at around 3000 Hz to altos, tenors and basses only (Doscher 1994), it appears here that for the lower pitches, this may be a formant tuning strategy for sopranos as well. The increased amplitude at the 3000 Hz frequency range in female voices when singing in the chest and middle register is also noted by Richard J. Morris and colleagues (2016). The results suggest that contrary to the general assumption that vowel modification only happens midway through a sung pitch range, the adjustment of vowel spaces to accommodate for voice projection occurs from the very beginning.



Comparison of spoken (Left) and sung (Right) [ta:˥] by participant four at around 250 Hz.

The results of the production task show how singers use vowel modification when singing through the pitch range. While this is expected to help sopranos achieve timbral coherence and voice projection, as the name vowel modification suggests, vowels may be sacrificed. In order to understand how vowel modification affects vowel intelligibility, a perception task was conducted. It was hypothesized that vowel perceptibility would reflect the formant analyses from the production task.

4 Experiment Two

4.1 *Methods and Materials* 79 native speakers of Cantonese, between 18 and 50 years old, with no vocal training in classical voice participated in the experiment. Participant one's recordings from the production experiment were used in the second experiment. Only a range of A4 (440 Hz) to G#5 (831 Hz) was included in the experiment to minimize the overall experiment length. Part of the spoken tokens were also used as stimuli for a screening task. Only [p] and [t] environments were considered to minimize the length of the experiment. [l] was excluded since it demonstrated similar coarticulatory effects to [t].

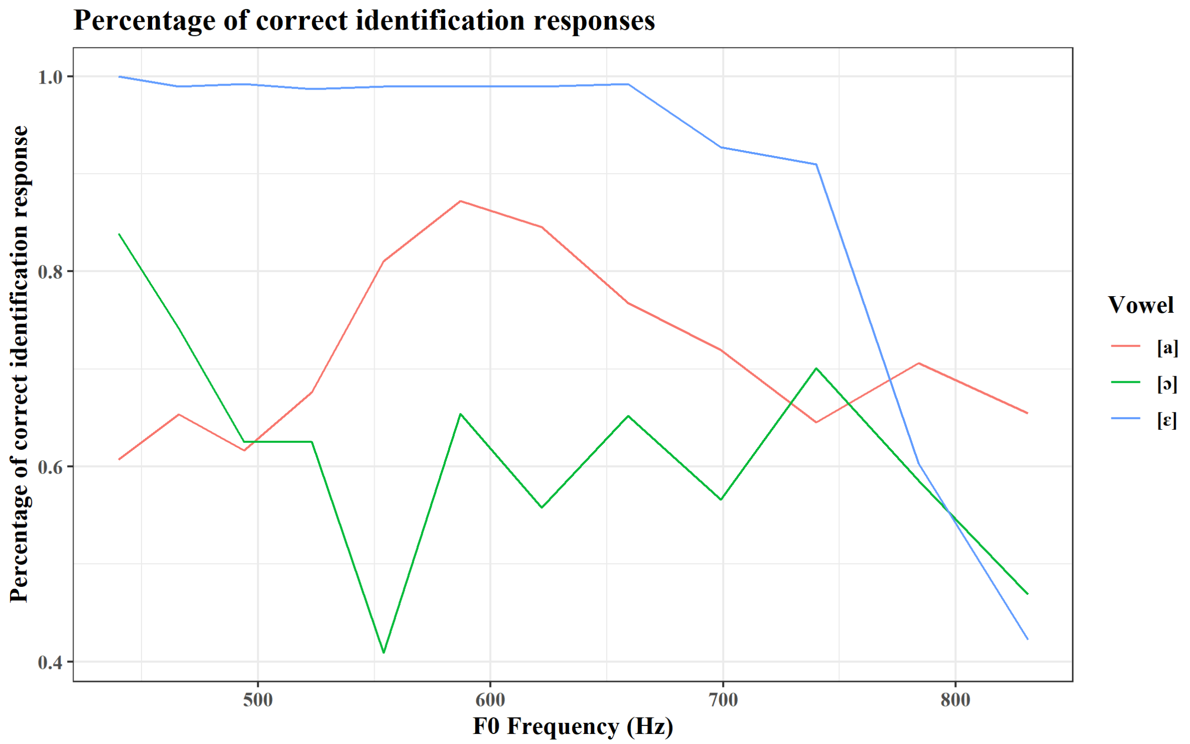
The stimuli were presented in the form of an AX discrimination task, while answer choices were presented as an identification task. This means that tokens were presented in pairs including the following combination of sounds: different vowels sung in the same pitch; same vowels sung in the same pitch; different vowels sung in different pitches; and same vowels sung in the same pitches. When stimuli had different F0s, they were always presented a whole tone apart. The onset consonant did not vary. All vowel permutations were included.

Participants were divided into two groups to minimize experiment length. Every group was exposed to all combinations of pitches and vowels but had alternating phonological environments. The order of token presentation was randomized. Each participant answered 9 questions that were speech based, those who passed the screening task answered an extra 99 questions that were based on the sung stimuli. The tokens were normalized by scaling the peak to 0.99, with a duration of 680 ms, equating to 88 bpm. The speech tokens were also normalized to a peak of 0.99 and a duration of 320 ms. A 500 ms gap was left in between the two tokens in each recording.

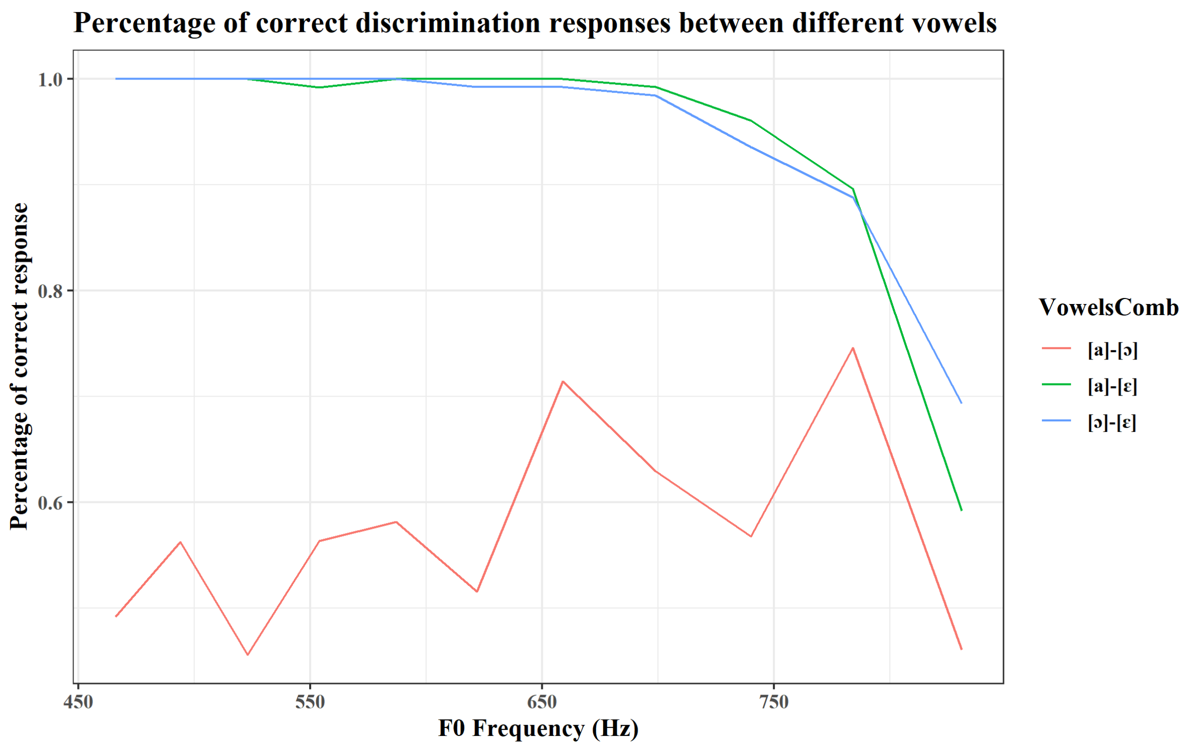
The experiment was conducted using an online survey software, Qualtrics. Participants were asked to complete the experiment in a quiet environment. Upon answering questions concerning their musical background and completion of the screening task, the participants who answered 7 out of 9 screening questions proceeded to the main task where they pressed the play button for each question at their own convenience. After listening to the two tokens in the question, they chose the word that they heard in the stimuli. All answer choices were presented in traditional Chinese characters. Participants were allowed to replay stimuli recordings if they wish. All questions were randomized.

The experiment was conducted as a mixture of identification and discrimination tasks. Based on the participants' identification response, their response in a discrimination task setting could be interpolated. Both results are important as it is conceivable that a participant hears [ta] and [ta] at a high pitch, when they are instead [tɔ] and [tɔ]. In that case, an incorrect vowel identification answer could still be considered correct in an AX discrimination task.

4.2 *Results* The results of the identification results are shown below, the identification rates of [a] and [ɔ] fluctuate greatly across all pitches. For the vowel [ɛ], identification rates dropped from around E5 (659 Hz) onwards. The overall results do not appear to support the hypothesis well.



Below shows the graph for the discrimination results. It shows that while [a]-[ɛ] and [ɔ]-[ɛ] pairs are distinguished well up to F5 (699 Hz), participants have difficulty discriminating between [a]-[ɔ] throughout the pitch range. The results roughly correspond to the identification results.



4.3 *Discussion* The hypothesis that there will be a gradual decline in vowel intelligibility according to formant change is not entirely supported. Instead, the back vowels tend to have high confusability at a pitch as low as 440 Hz. It is worth speculating why [ɛ] seems to behave differently compared to [a] and [ɔ]. Potential phonological factors that may have influenced the vowels' behaviors may include the relative position of the vowels and the distance between the vowels. Potential phonetic reasons would include how much head tone the soprano used in her mixed register, since the strengthening of harmonics may contribute to vowel perceptibility

(Gottfried & Chew 1986). Either way, the results show that when sung vowels are concerned, there may be different strategies to vowel modification, and vowel intelligibility should also be considered on a vowel-by-vowel basis.

5 Overall Discussion

From the results of the perception task, it seems that F2 may be a better predictor for vowel intelligibility compared to F1. Considering [a] and [ɔ], both their F1s and F2s are considerably close by, reflecting its overall confusability. However, the F1 of [ɛ] in the production task does seem to start modifying in a more stable manner at around E5 (659 Hz), while the F2 drop has been continuing regardless of pitch change, suggesting that F1 could be an important factor to [ɛ]'s decreasing intelligibility. The role of F2 remains inconclusive. Statistical analyses of the production task using a mixed model may be used to formalize these observations and comparisons better in the future.

6 Conclusion

The results of the present study show that in studying vowel modification, individual differences and gender differences should be taken into account. Although vocal pedagogy may describe vowel modification as one uniform technique across all voice types, in practice, the nuances are often not described empirically. The present study also shows that vowels may behave differently during vowel modification as well, which results in varying degrees of vowel intelligibility even within the same pitch range. Vowel formants are also shown to potentially take on non-linguistic roles, holding implications to vocal pedagogy and our understanding of voice projection and timbral control.

Research gaps to consider include how registers and the *passaggio* should be defined acoustically. Existing studies often consider single vowels and laryngeal activity, disregarding the differences between vowels and the role of singers' resonance space (Morris et al. 2016, Miller & Schutte 2005). The present study suggests that different vowels may require different approaches when singing through the *passaggio,* and that vowel formants may take on non-linguistic rules in lower pitches as well. The present study shows that when studying the singing voice, both its musical aspects and the language's phonology should be considered, calling for closer connection between musicological and linguistic studies.

References

Callaghan, Jean. 2000. *Singing and voice science.* San Diego, CA: Singular Publishing Group.

Campbell, Murray. 2001. Timbre. *Grove Music Online.* Online: https://doi.org/10.1093/gmo/9781561592630.article.27973.

Doscher, Barbara. 1994. *The Functional Unity of the Singing Voice*. Metuchen, N.J.: Scarecrow Press.

Gottfried, Terry L, and Stephen L. Chew. 1986. Intelligibility of vowels sung by a countertenor. *The Journal of the Acoustical Society of America* 79(1).124-130.

International Phonetic Association. 1999. *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

Kwan, Tze-wan. 2014. *Multi-function Chinese Character Database*, 8 July 2014. Online: http://humanum.arts.cuhk.edu.hk/Lexis/lexi-mf/

Ladefoged, Peter, and Johnson, Keith. 2010. *A Course in Phonetics.* Boston, Wadsworth: Cengage Learning.

Lewcock, Ronald; Rijn Pirn, Jürgen Meyer, Carleen M. Hutchins, J. Woodhouse, John C. Schelleng, Bernard Richardson, Daniel W. Martin, Arthur H. Benade, Murray Campbell, Thomas D. Rossing and Johan Sundberg. 2001. Acoustics. *Grove Music Online.* Online: https://doi.org/10.1093/gmo/9781561592630.article.00134

Miller, Donald G.. 2000. R*egisters in Singing: Empirical and Systematic Studies in the Theory of the Singing Voice.* Groningen: Rijksuniversiteit Groningen.

Miller, Donald G., and Harm K. Schutte. 2005. 'Mixing' the Registers: Glottal Source or Vocal Tract?. *Folia Phoniatrica Et Logopaedica* 57(5-6).278-291.

Miller, Richard. 2000. *Training Soprano Voices.* Oxford: Oxford University Press.

Morris, Richard J.; David Okerlund; and Emily A. Craven. 2016. First Passaggio Transition Gestures in Classically Trained Female Singers. *Journal of Voice* 30(3).377.e21-377.e29.

Rothenberg, Martin. 1984. Source-tract acoustic interaction and voice quality. *Transcripts of the Twelfth Symposium: Care of the Professional Voice, 1983,* ed. by Van L. Lawrence, 25-31. New York: The Voice Foundation.

Sundberg, Johan. 1999. The perception of singing. *Academic Press series in cognition and perception: A series of monographs and treatises. The psychology of music,* ed. by Diana Deutsch, 171-214. San Diego, CA: Academic Press.